THE PERCEPTION AND PREFERENCE OF COMPOUND SPACES IN TRADITIONAL VILLAGES IN CHINA (1066)

Jia Feng[1*], Xiaojian Chen[1]

[1] School of Architecture, Xi'an University of Architecture and Technology, China; [*]fengjia91@163.com

Abstract. Traditional villages in China are undergoing spatial renewal and reorganization, resulting in the emergence of compound spaces comprising modern and historical buildings. As integral parts of these traditional villages, these compound spaces have a significant impact on the perception preferences of the historical spaces within them. This study employs deep learning techniques to analyze photos and videos containing compound spaces and extract the visual occupation ratio and visual change ratio of modern and historic buildings. The specific focus is to explore the influence of the visual occupation ratio and visual change ratio on spatial perception preferences, aiming to analyze and understand the delicate balance between sustainable development and heritage preservation in traditional villages. The paper examines the external representation of architectural elements within compound spaces, including the potential increase in perception preferences by reducing modern buildings and the influence of trends in modern and historical buildings on perception preferences. The findings from this research provide valuable insights for addressing spatial dissonance amidst rapid development in traditional villages in China.

Keywords: Perception, Preference, Visual Occupation Ratio, Compound Spaces, Traditional Villages

1. Introduction

Traditional villages are an important part of China's rural heritage, embodying the country's rich farming civilization. They hold historical, cultural, and ecological significance and serve as repositories of historical memory(Liu & Xu, 2021). Recognizing their value, the Chinese government has introduced laws and policies since the beginning of the new millennium to protect and revitalize traditional villages(Guo & Sun, 2016). However, the rapid renewal and reorganization of these villages have led to disorderly development and construction. This haphazard expansion has negatively impacted the spatial features of the villages(Zhao et al., 2016) undermining their historical charm and endangering their heritage and cultural value.

Early research on the perception of traditional village aesthetics primarily focused on landscape feature analysis, evaluation and optimization, as well as protection and management. The objective was to analyze the composition, development, evolution, essential characteristics, and existing issues of village spaces. The research was divided into two approaches: Imagism and Imageism. Imagism drew upon the research methodology of Kevin Lynch's urban image theory, utilizing spatial intention maps from different groups within the village to explore the deep-level spatial continuation in the process of village evolution. This approach integrated space syntax analysis to identify invariant elements within the spatial style. On the other hand, Imageism employed picture data as the research object, drawing upon methods such as psychological scales, emotional evaluation techniques, cognitive scorecards, visual preference surveys, and scenic beauty evaluation methods. Experimental methods were used to understand the relationship between village aesthetics and people's aesthetic responses, while assessing variations in individual participants' impressions of the village's external appearance(Foltête et al., 2020)(Jiang et al., 2015)(Zhang & Lin, 2011)(Gobster et al., 2019). During this period, the image data aimed to reveal perceptual preferences related to large-scale landscape elements such as heritage environment, agricultural environment, and natural environment(Yang et al., 2021).

In recent years, significant breakthroughs have been achieved in using computational and data techniques to analyze the feature information contained in images through deep learning (DL) algorithms. Compared to traditional machine learning approaches, intelligent recognition based on deep learning and large-scale image datasets has greatly improved accuracy. Within the research paradigm of deep learning, each image possesses distinctive features that set it apart from others. Some of these features are natural and perceptible, such as brightness, edges, textures, and colors, while others may require transformation or processing to be observed. This research approach aids in analyzing the composition and characteristics of physical environmental elements in villages. For instance, by identifying traditional village elements and socio-environmental attributes in images and utilizing image recognition based on physical environmental features(Yin & Wang, 2016), the perceptual preferences of people's cognitive concepts can be linked, enabling intelligent evaluations of village style (Cheng et al., 2017) and visual quality(L. Liu et al., 2017). Currently, deep learning techniques primarily focus on image recognition and evaluation, rural landscape identification, evaluation, and protection in the research on the perception of traditional villages.

However, existing research on traditional villages has primarily focused on the overall spatial characteristics, with limited exploration of the perceptual preferences regarding compound spaces formed by the combination of modern and historical architecture. For instance, there is a lack of study on the impact mechanism of the combination of modern and historical architecture on people's perceptual preferences and which tendencies in modern and historical architecture may increase or decrease such preferences. This study employs deep learning methods to extract visual occupancy and visual change rates of modern and historical architecture from photographic and video data containing compound spaces. In conjunction with neurophysiological experiments, this research investigates people's perception and preferences regarding the compound spaces in traditional villages.

2.Methods

2.1 Data Collection And Preprocessing

2.1.1 Data Collection Method

The collection of visual data was accomplished using the Gopro8 model from Gopro company. The video functionality of this device allowed for continuous and intermittent recording of image data and coordinate positions based on the user's travel distance. Additionally, the device provided various settings and modes, such as timed burst mode and video mode, which facilitated the use of different options when capturing environmental images. By connecting the device with the accompanying application on a smartphone, the surveyors were able to traverse predefined routes. The built-in stabilization feature of Gopro effectively reduced errors between the recorded video data and the actual environment caused by shaking and vibration. For this study, the entire environmental images were captured in video mode. All collected data were saved in a database.

2.1.2 Survey Location

In order to minimize the potential errors resulting from regional variations in geographical, economic, and cultural environments, we selected five villages within a concentrated and contiguous protection area of traditional villages as the experimental subjects. This area is located in Jinhua City, Zhejiang Province, China. The data collection process commenced at the same time over a span of five working days. We predetermined the areas to be measured, and the surveyors collected data by walking through the villages. Based on the proportional relationship between historical and modern buildings in traditional villages, the entire environmental data were categorized into six types: Single TypeⅠ(a), Single TypeⅡ(b), Mixed TypeⅠ(c), Mixed TypeⅡ(d), Mixed TypeⅢ(e), and Natural Type(f) (Figure 1).

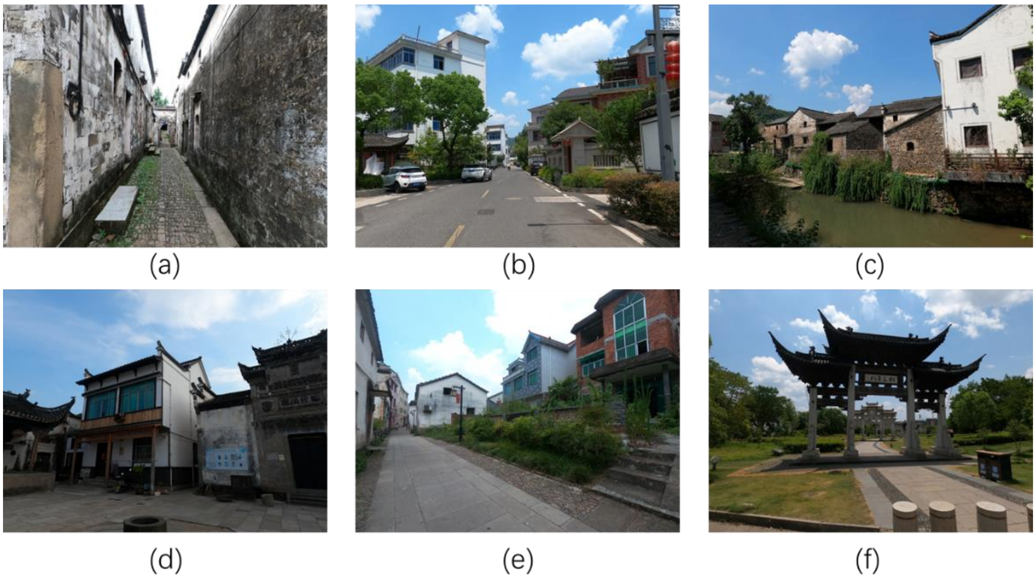


Figure 1. The six different proportional types of composite spaces in traditional villages are represented as follows: (a) historical architectural spaces, (b) modern architectural spaces, (c)historical spaces with low hybridity, (d)complex spaces with high hybridity, (e)modern spaces with low hybridity, (f)natural spaces.

The collected data underwent preprocessing to select video data that effectively reflected the environmental characteristics. In the preprocessing stage, video data was segmented into 1-minute clips(Li & Kang, 2019) from the original video data. Finally, through the preprocessing step, we identified and selected 20 video data clips.

2.2 Data Analysis

2.2.1 Deep Learning Model

Deep learning (DL) is a method of learning the intrinsic patterns and hierarchical representations of sample data by extracting indicators from collected image datasets. The model used in this study is Convolutional Neural Network (CNN), which is primarily employed in computer vision tasks such as semantic segmentation (SS). DL requires image datasets for training, and currently available open datasets include ADE20K, ImageNet, and others, which contain basic environmental elements of cities but fail to differentiate between historical and modern architecture. In this research, we developed a small-scale semantic annotation database specifically for traditional village images. In addition to the dataset, we utilized MobileNetV2 and PSPNet as the architecture for the entire deep learning model, with its primary function being video semantic segmentation (VSS) (Figure 2). The visual dataset training was conducted using PyTorch, and the training and utilization of the entire model were performed in PyCharm.

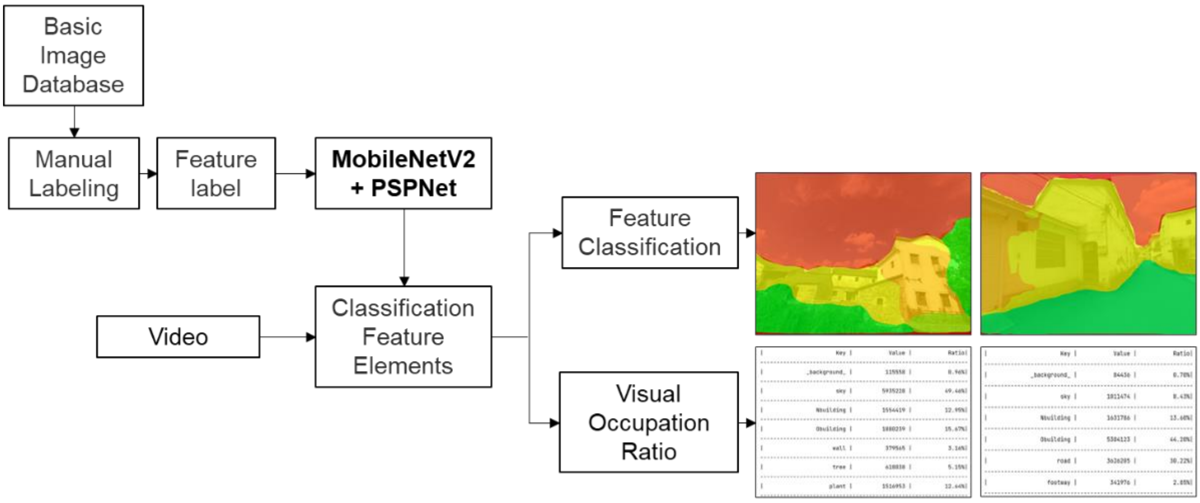


Figure 2. Demonstrates the workflow and results of the semantic segmentation task (VSS) based on MobileNetV2+PSPNet

2.2.2 Measure From Videos

To quantify the visual occupancy rates(VOR) of historical and modern architecture in the videos, we employed the VSS function of the SS model to measure the percentage of pixels occupied by modern and historical architecture per second in each video. Pixels are the smallest units that can be displayed and controlled in a raster image. Digital photographs consist of a two-dimensional grid of pixels. By inputting photos containing historical and modern architecture into the SS model, we obtained sets of percentages representing the pixel occupancy of modern architecture (N-VOR) and historical architecture (O-VOR) in the video data on a per-second basis. Based on the acquired VOR features, we defined different spatial types using three levels: high (H; VOR > 40%), moderate (M; 40% ≥ VOR > 10%), and low (L; 10% ≥ VOR). To distinguish between historical and modern architecture, we used Ho, Mo, Lo, and Hn, Mn, Ln to represent different levels of O-VOR and N-VOR, respectively.

2.3 Physiological Measurements

2.3.1 Experimental Procedure

We employed an experimental approach to measure the physiological perception process of the video data. Physiological indicators have been proven to be more suitable for studying continuous data. The experiments were conducted in a university classroom with controlled room temperature (22°C), air humidity (65%), and illumination (500lx). There were no other distracting objects in the room. Participants were instructed to sit comfortably at a distance of 1.5m from the screen. The experiments were conducted throughout the day, from 9:00 AM to 11:30 AM and from 2:00 PM to 5:00 PM. The study aimed to provide the same environmental conditions for each category to reduce errors. The procedure for each participant was as follows: the experimental process and requirements were explained to the participants. After obtaining their consent, the researchers attached the electroencephalogram (EEG) devices to the participants. Calibration of the baseline EEG data was performed first. The videos were also processed to remove sound to minimize its impact on the experiment. Subsequently, the participants randomly selected a type of video data, and EEG data were recorded during this period (Figure 3). A total of 398 EEG data recordings were collected in this experiment.

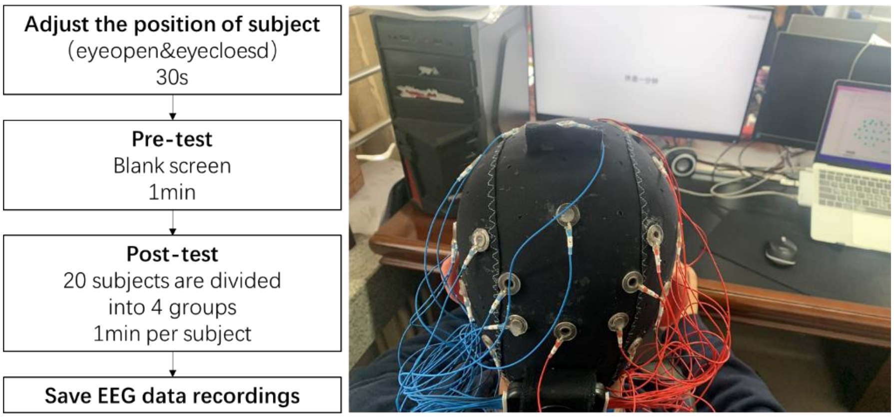


Figure 3. Experimental procedure

2.3.2 Physiological Measurements

This experiment utilized the Emotiv EPOC FLEX system to detect physiological signals. It is a non-invasive and harmless electroencephalogram (EEG) device. The reliability and accuracy of the Emotiv EEG device have been validated in previous studies. The device consists of 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4), covering four regions of the brain (frontal, temporal, parietal, and occipital lobes). The EPOC FLEX system continuously captures the brainwave signals from the entire brain (i.e., 14 electrodes) and transmits these signals to a computer hard drive via Bluetooth for storage as raw EEG data. The Emotiv Pro software calculates the participants' performance metrics based on the raw

EEG data, including Engagement, Excitement, Focus, Interest, Relaxation, and Stress dimensions. The device generates data every 10 seconds.

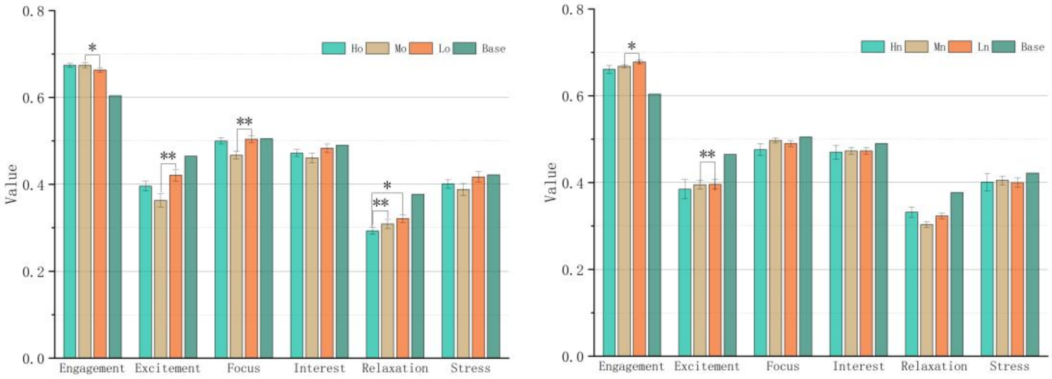
2.4 Data Analysis

The statistical analyses were conducted using SPSS version 25. Firstly, the mean values of the Visual Occupancy Rate (VOR) for historical and modern architecture were calculated at a 10-second interval. Additionally, the Visual Change Rate (VCR) was defined based on the changes in VOR between consecutive 10-second intervals. Covariance analysis (ANCOVAs) was performed with the resting-state physiological indicators as covariates and O-VOR and N-VOR as independent variables to analyze the main effects and interaction effects. Subsequently, simple effects analyses were conducted on the physiological indicators showing significant interactions. Furthermore, the relationship between spatial VCR and changes in physiological indicators was statistically analyzed using k-means cluster analysis.

3. Results

3.1 Main Effect Analysis Of VOR And Physiological Indicators

We analyzed the main effects of O-VOR and N-VOR on six physiological indicators: engagement, excitement, focus, interest, relaxation, and stress (Figure 4, Table 1). The results



indicated significant correlations between baseline physiological indicators and the outcomes. The main effects of engagement and interest were not significant, and there was no significant interaction effect. Excitement measured before and after the experiment showed significant differences, with a decrease in excitement during the experimental process. The main effect of O-VOR was significant, with the following corresponding results: Lo (M=0.421) > Ho (0.396) > Mo (0.363). O-VOR also showed a significant difference in its impact on focus, with the following results: Lo (M=0.504) > Ho (M=0.500) > Mo (M=0.467). Additionally, the results demonstrated a significant interaction effect between O-VOR and N-VOR on focus. There were significant differences in the effects of O-VOR and N-VOR on relaxation. The main effect of O-VOR was significant, with the following results: Lo (M=0.321) > Mo (M=0.309) > Ho (0.293). The main effect of N-VOR was significant, with the following results: Hn (0.332) > Ln (0.323) > Mn (0.303). The stress values decreased before and after the experiment, but the differences in the effects of O-VOR and N-VOR on the stress indicator were not significant, although the interaction effect was significant.

Figure 4. Value of Physiological indicators in the main effect analysis(N=1155,standard error, * $p < 0.05$； ** $p < 0.01$).

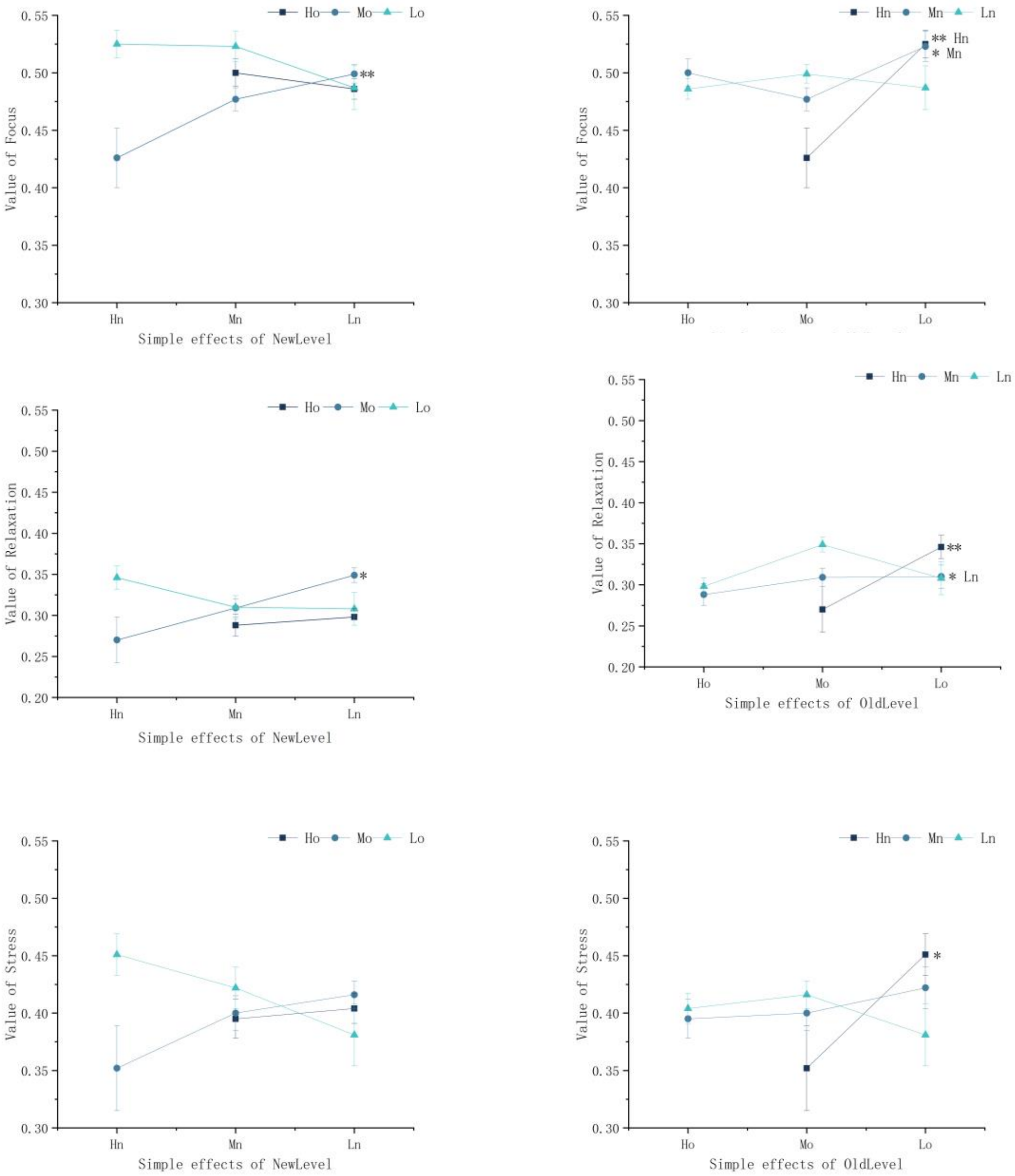
3.2 Simple Effect Analysis Of Significant Interaction Effects

The results of the main effect analysis showed significant differences in the values of the Fcous, Relaxation, and Stress physiological indicators before and after the experiment, showing a decreasing trend (Figure 5, Table 2). The interaction analysis between N-VOR and O-VOR revealed that these three indicators were not only influenced by the main effects but also by the interaction effects, necessitating further analysis through simple effect analysis to explain the different roles of N-VOR and O-VOR.

The simple effect results for Fcous are as follows: Mo-Ln (M=0.499) > Mo-Hn (M=0.486); Hn-

Lo (M=0.525) > Hn-Mo (M=0.426); Mn-Lo (M=0.523) > Mn-Mo (M=0.477). The simple effect analysis further demonstrated that the interaction effect was stronger for Mo-Ln compared to Mo-Hn. The interaction effect of Hn-Lo was greater than Hn-Mo, and the interaction effect of Mn-Lo was also greater than Mn-Mo. In spaces where one element has a low visual occupancy rate, the combined effect of H-VOR and N-VOR is higher on Fcous. In other words, in compound spaces with lower mixing rates, the visual occupancy rate of the buildings has a positive impact on Fcous, and lower mixing rates are more likely to increase Fcous neurons.

The simple effect results for Relaxation are as follows: Mo-Ln (M=0.349) > Mo-Hn (M=0.270),



Mo-Ln (M=0.349) > Mo-Mn (M=0.309); Ln-Lo (M=0.308) > Ln-Ho (M=0.298). The simple effect analysis further demonstrated that the interaction effect of Mo-Ln was greater than Mo-Hn and Mo-Mn, indicating that in historical spaces, the lower the proportion of modern spaces, the higher the Relaxation. The interaction effect of Ln-Lo was greater than Ln-Ho. This result suggests that in lower levels of N-VOR, a smaller proportion of O-VOR can increase people's Relaxation.

Figure 5. Value of Physiological indicators in the simple effect analysis(N=1155,standard error, * $p < 0.05$； ** $p < 0.01$).

The simple effect results for Stress are as follows: Hn-Lo (M=0.451) > Hn-Mo (M=0.352). The simple effect analysis further demonstrated that in higher levels of N-VOR, a smaller proportion of O-VOR leads to lower Stress values. In other words, when a large number of historical buildings are reconstructed as modern buildings, the Stress value of the space increases."

3.3 Cluster Analysis Of Visual Change Rates

In addition to the VOR analysis, we conducted a cluster analysis of visual change rates(VCR) in relation to changes in the six physiological indicators. The results of the study showed that when there were changes in VCR within the space, corresponding changes were observed in the six physiological indicators among participants. We statistically analyzed the differences between post-test and baseline physiological indicators (pos-pre) as a measure of VCR changes (Figure 6). We further performed K-means cluster analysis on the change features, with the number of clusters set to 5 (Table 3). Cluster 1, Cluster 2, Cluster 3, Cluster 4, and Cluster 5 represented negative-strong change, negative-moderate change, weak change, positive-moderate change, and positive-strong change, respectively.

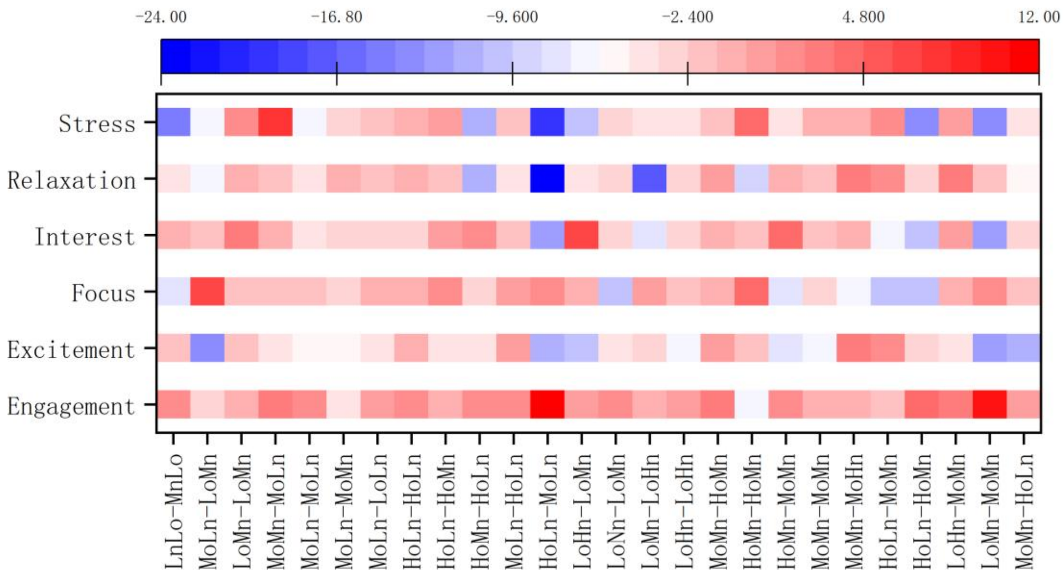


Figure 6. Statistical Heatmap of Physiological Indicator Changes due to Different VCRs

The results revealed that LnLo-MnLo indicated a decrease in stress when transitioning from open spaces to historical spaces at a moderate level. MoLn-LoMn indicated an increase in focus and a decrease in excitement when transitioning from historical spaces with moderate VOR to modern spaces with moderate VOR. The results of MoMn-MoLn suggested that transitioning from spaces with high mixing rates to spaces with low mixing rates increased both stress and relaxation. HoLn-MoLn represented an increase in engagement and a decrease in stress and relaxation when transitioning from high-density historical spaces to moderate-density historical spaces. The results of LoHn-LoMo and LoMn-LoHn indicated that reducing the proportion of modern buildings in a space primarily composed of modern architecture increased interest and relaxation. LoMn-MoMn represented an increase in engagement and a decrease in interest when transitioning from spaces with moderate-density modern buildings to mixed spaces.

4. Discussion

There is a growing body of research on the perception preference of heritage spaces, and most studies have discussed the impact of aesthetic qualities of heritage landscapes on people's well-being(Sektani et al., 2021). For example, exploring the influence of historical urban environment features on public visual preferences(Deghati Najd et al., 2015). Many studies have assessed the effects of perceived preferences for different types of heritage spaces in urban environments. However, only a few studies have focused on the differences between heritage physical buffer zones and perceptual buffer zones, and there is limited research on the proportion of historical and modern buildings in combination. Furthermore, there is a lack of research on the neural and physiological mechanisms underlying the relationship between visual occupancy rates and perception from a neurophysiological perspective. Importantly, there have been no reports on the dynamic perception preference of heritage spaces from a neurophysiological perspective. Our study found that O-VOR and N-VOR have varying degrees of influence on the experience of traditional village compound spaces.

4.1 Continuous Historical Spaces Are More Likely To Enhance Focus And Excitement

The diversity of landscapes experienced by individuals is limited to appropriate patterns, while landscape clutter refers to an increase in landscape diversity accompanied by the invasion of other visual elements and a lack of coherence, which affects our perception of the environment(Wagtendonk & Vermaat, 2014). Our study found a significant main effect of O-VOR on focus and excitement, while N-VOR did not have a significant main effect, demonstrating that historical spaces can control focus and excitement. Visually, high O-VOR and low N-VOR share a common characteristic of high continuity of elements. In low O-VOR spaces, the sky and greenery exhibit high continuity, while in high O-VOR spaces, historical buildings exhibit high continuity. However, in spaces with medium O-VOR, various elements are mixed together, especially the interaction between N-OVR and O-VOR, making it difficult for individuals to sustain their attention on a particular element. Additionally, the interaction between O-VOR and N-VOR significantly affects focus, indicating that N-OVR has an influence on the continuity of historical spaces to some extent. This is manifested by the fact that in spaces with medium to high O-VOR, fewer NVOR elements result in less disruption to the continuity of historical spaces.

4.2 Low Mixing Ratio Of O-VOR And N-VOR Can Enhance Relaxation

Furthermore, we found significant main and interaction effects of O-VOR and N-VOR on relaxation, indicating that these two factors not only individually influence relaxation but also interact with each other to affect relaxation. Analyzing the results, O-VOR and N-VOR yield completely opposite outcomes. O-VOR shows a negative correlation with relaxation, which may be attributed to the fact that high O-VOR typically represents narrower street and alley spaces, resulting in a higher aspect ratio that increases environmental stress and reduces relaxation. The characteristics of the N-VOR index also indirectly support this conclusion. In traditional villages, Ln and Mn represent Ho, while Hn spaces often have a minimal or non-existent presence of O-VOR. Therefore, Hn has a more positive impact on relaxation compared to Mn and Ln. Regarding the significance of the interaction effect, in historical spaces, a lower proportion of modern spaces generates higher relaxation. In lower levels of N-VOR, a smaller proportion of O-VOR increases relaxation. Previous studies have focused on the negative effects of modern buildings surrounding historical heritage. This study further demonstrates, from a neurophysiological perspective, that a low mixing ratio of O-VOR and N-VOR enhances spatial relaxation. The aforementioned results confirm that monotonous environments are more conducive to the restoration of focused attention and alleviation of mental fatigue, leading to relaxation compared to complex environments. Additionally, the results of simple effect analysis demonstrate that individuals are more likely to experience higher relaxation in environments dominated by O-VOR, with the most pronounced positive effect observed in Mo's O-VOR.

4.3 Significant Increase In N-VOR In Heritage Spaces Can Elevate Stress

Visual perception plays a dominant role in human-environment interaction. Previous studies have confirmed the positive impact of aesthetic sensory experiences on well-being(Sektani et al., 2021), where buildings with perceived aesthetic quality positively influence individuals' happiness, while unattractive buildings have detrimental effects. Our study found that the main effects of O-VOR and N-VOR on stress were not significant, but the interaction effect was significant. The possible reason is that the individual indicators of historical and modern buildings do not directly influence spatial stress, but their mutual variations affect spatial stress. From the results of the simple effect analysis on stress, the reconstruction of historical buildings into modern ones significantly increases spatial stress. One possible reason is that in traditional villages, haphazardly constructed modern buildings are considered unattractive and have a detrimental impact on our sense of well-being, thus generating higher stress levels.

4.4 The Impact Of Spatial VCRs Changes On Physiological Indicators

In this study, the influence of entering different types of spaces on individual experiences was analyzed. The results demonstrate a certain relationship between different spatial characteristics and individuals' psychological states. Firstly, the findings indicate that transitioning from open spaces to moderately historical spaces reduces the sense of stress. This could be attributed to open spaces providing better comfort and relaxation, while historical spaces exert a positive influence due to their unique cultural and historical value. Secondly, the results show that transitioning from moderate VOR (historical spaces) to moderate VOR (modern spaces) increases focus and decreases excitement. This may be because individuals find it easier to concentrate in historical spaces and experience fewer stimuli in modern spaces, leading to reduced excitement. Furthermore, the results suggest that transitioning from highly mixed spaces to low-mixed spaces increases individuals' stress and relaxation. This could be due to the need for individuals to adapt to diverse environments in highly mixed spaces, while low-mixed spaces result in reduced environmental stimuli, thereby increasing stress and relaxation. Additionally, the results demonstrate that transitioning from high-density historical spaces to moderate-density historical spaces enhances individuals' sense of engagement and reduces stress and relaxation. This may be because high-density historical spaces make individuals more aware of the compactness of the environment, thus increasing their sense of engagement. Finally, the results indicate that in spaces predominantly characterized by modern architecture, reducing the proportion of modern buildings increases individuals' interest and relaxation. This may be because modern buildings tend to be more stimulating, and reducing the proportion of modern architecture can create a more attractive and relaxing environment.

5.Conclusion

This study explores the effects of different combinations of O-VOR and N-VOR in complex traditional village spaces on individual neuro-psychological responses. By investigating people's perceptual preferences for different types of heritage spaces, we examined the influence of the landscape aesthetic quality of heritage spaces on individual happiness. The results demonstrate the positive impact of continuous historical spaces on enhancing attention and excitement. Historical spaces with higher physical continuity and lower perceived clutter contribute to individuals focusing their attention on specific elements and experiencing greater excitement. Conversely, excessive landscape clutter and lack of coherence have a negative impact on individuals' environmental perception. Furthermore, a lower proportion of O-VOR and N-VOR with low mixing ratios can enhance individuals' relaxation. Compared to higher proportions of heritage physical buffer zones, lower combinations create a simpler, more monotonous environment that helps alleviate mental fatigue and enhance relaxation. This suggests that in heritage spaces, an excessive presence of modern buildings may increase individuals' stress and tension, while fewer modern buildings can provide a better relaxation experience. However, the results also indicate that a significant increase in modern buildings within heritage spaces may increase individuals' sense of stress. In traditional villages, haphazardly constructed modern buildings, considered unattractive structures, may negatively impact individuals' sense of happiness, resulting in higher levels of stress. In summary, this study provides a comprehensive understanding of the neuro-psychological effects of different combinations of O-VOR and N-VOR in complex traditional village spaces. These findings have important implications for urban planning and heritage preservation, offering guidance for creating pleasant and relaxing environments in traditional villages.

Further research can explore the perceptual preferences for heritage spaces among different populations and differences across various cultural backgrounds. Additionally, studying the influence of different age groups, genders, cultural backgrounds, and socioeconomic conditions on the perception of heritage spaces can provide insights. Understanding these differences can help us better meet the needs of diverse populations and design heritage spaces that are inclusive and diverse. Finally, the findings of this study hold significance for urban planning and heritage preservation practices. We should prioritize the protection and restoration of heritage spaces while considering the psychological needs of individuals. Through thoughtful planning and design, we can create heritage spaces that have positive impacts, offering enjoyable, relaxing, and meaningful experiences for people. In conclusion, this study provides valuable insights into the neuro-psychological effects of different combinations of O-VOR and N-VOR in complex traditional village spaces, shedding light on the relationship between spatial characteristics and individual experiences. These findings offer useful guidance for creating more attractive, comfortable, and satisfying heritage spaces.

Foundation: The National Key Research and Development Program of China (Project No. 2019YFD1100905)

Appendix

Table 1. Main effects and interaction effects of O-VOR and N-VOR on the physiological indicators across the experimental groups (covariate: pre-test; dependent variable: post-test).

Parameters	Sum of Squares	df	Mean Square	F	Sig.	Partial η2	Pairwise Comparisons
DV：Engagement							
O-VOR	0.012	2.000	0.006	0.783	0.457	0.001	
N-VOR	0.029	2.000	0.015	1.900	0.150	0.003	
O-VOR * N-VOR	0.002	3.000	0.001	0.092	0.965	0.000	
Engagement_Base	0.609	1.000	0.609	79.138	0.000**	0.065	
Error	8.823	1146.000	0.008				
R^2 = 0.078（Adj. R^2 = 0.071）							
DV：Excitement							
O-VOR	0.412	2.000	0.206	4.554	0.011*	0.008	Lo > Ho > Mo
N-VOR	0.009	2.000	0.004	0.098	0.907	0.000	
O-VOR * N-VOR	0.180	3.000	0.060	1.331	0.263	0.003	
Excitement_Base	6.056	1.000	6.056	133.998	0.000**	0.105	
Error	51.794	1146.000	0.045				
R^2 = 0.121（Adj. R^2 = 0.115）							
DV：Focus							
O-VOR	0.202	2.000	0.101	5.803	0.003**	0.010	Lo > Ho > Mo
N-VOR	0.030	2.000	0.015	0.848	0.429	0.001	
O-VOR * N-VOR	0.184	3.000	0.061	3.515	0.015*	0.009	
Focus_Base	1.869	1.000	1.869	107.166	0.000**	0.086	
Error	19.986	1146.000	0.017				
R^2 = 0.103（Adj. R^2 = 0.097）							
DV：Interest							
O-VOR	0.055	2.000	0.028	1.133	0.322	0.002	

N-VOR	0.001	2.000	0.001	0.022	0.978	0.000	
O-VOR * N-VOR	0.149	3.000	0.050	2.031	0.108	0.005	
Interest_Base	3.264	1.000	3.264	133.567	0.000**	0.104	
Error	28.005	1146.000	0.024				
R^2 = 0.112（Adj. R^2 = 0.106）							
DV：Relaxation							
O-VOR	0.137	2.000	0.069	3.002	0.049*	0.005	Lo > Mo > Ho
N-VOR	0.141	2.000	0.071	3.039	0.048*	0.005	Hn > Ln > Mn
O-VOR * N-VOR	0.244	3.000	0.081	3.552	0.014*	0.009	
Relaxation_Base	2.394	1.000	2.394	104.552	0.000**	0.084	
Error	26.242	1146.000	0.023				
R^2 = 0.103（Adj. R^2 = 0.096）							
DV：Stress							
O-VOR	0.091	2.000	0.046	1.174	0.310	0.002	
N-VOR	0.005	2.000	0.003	0.066	0.936	0.000	
O-VOR * N-VOR	0.289	3.000	0.096	2.476	0.060*	0.006	
Stress_Base	0.560	1.000	0.560	14.413	0.000**	0.012	
Error	44.862	1151.000	0.039				
R^2 = 0.022（Adj. R^2 = 0.015）							

Note: O-VOR refers to the Visual Occupancy Rate of historical buildings. N-VOR refers to the Visual Occupancy Rate of modern buildings. Ho and Hn represent high visual occupancy rates for historical and modern buildings (above 40%). Mo and Mn represent medium visual occupancy rates for historical and modern buildings (40%-10%). Lo and Ln represent low visual occupancy rates for historical and modern buildings (below 10%). * $p < 0.05$; ** $p < 0.01$.

Table 2. Simple effects analysis of the parameters that have significant interaction effects (O-VOR × N-VOR) and significant pre-test and post-test differences.

Parameters	**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial η2**	**Pairwise Comparisons**

DV: Focus							
N-VOR(O-VOR=Ho)	0.016	1	0.016	0.845	0.358	0.001	
Error	21.855	1147	0.019				
N-VOR(O-VOR=Mo)	0.167	2	0.084	4.386	0.013*	0.008	Ln > Hn
Error	21.855	1147	0.019				
N-VOR(O-VOR=Lo)	0.063	2	0.031	1.653	0.192	0.003	
Error	21.855	1147	0.019				
O-VOR (N-VOR =Hn)	0.233	1	0.233	12.239	0.000**	0.011	Lo > Mo
Error	21.855	1147	0.019				
O-VOR (N-VOR =Mn)	0.152	2	0.076	3.977	0.019*	0.007	Lo > Mo
Error	21.855	1147	0.019				
O-VOR (N-VOR =Ln)	0.024	2	0.012	0.628	0.534	0.001	
Error	21.855	1147	0.019				
DV: Relaxation							
N-VOR(O-VOR =Ho)	0.008	1	0.008	0.346	0.556	0	
Error	26.242	1146	0.023				
N-VOR (O-VOR =Mo)	0.28	2	0.14	6.116	0.002**	0.011	Ln > Hn , Ln > Mn
Error	26.242	1146	0.023				
N-VOR (O-VOR =Lo)	0.096	2	0.048	2.104	0.122	0.004	
Error	26.242	1146	0.023				
O-VOR (N-VOR=Hn)	0.135	1	0.135	5.908	0.055	0.005	
Error	26.242	1146	0.023				
O-VOR (N-VOR =Mn)	0.043	2	0.022	0.948	0.388	0.002	
Error	26.242	1146	0.023				
O-VOR(N-VOR =Ln)	0.355	2	0.177	7.745	0.000**	0.013	Lo > Ho
误差	26.242	1146	0.023				

DV: **Stress**							
N-VOR(O-VOR =Ho)	0.006	1	0.006	0.151	0.697	0	
Error	45.097	1147	0.039				
N-VOR (O-VOR =Mo)	0.116	2	0.058	1.469	0.231	0.003	
Error	45.097	1147	0.039				
N-VOR (O-VOR =Lo)	0.189	2	0.094	2.397	0.091	0.004	
Error	45.097	1147	0.039				
O-VOR(N-VOR =Hn)	0.229	1	0.229	5.82	0.016*	0.005	Lo > Mo
Error	45.097	1147	0.039				
O-VOR (N-VOR =Mn)	0.051	2	0.025	0.644	0.525	0.001	
Error	45.097	1147	0.039				
O-VOR (N-VOR =Ln)	0.06	2	0.03	0.766	0.465	0.001	
Error	45.097	1147	0.039				

Note: O-VOR refers to the Visual Occupancy Rate of historical buildings. N-VOR refers to the Visual Occupancy Rate of modern buildings. Ho and Hn represent high visual occupancy rates for historical and modern buildings (above 40%). Mo and Mn represent medium visual occupancy rates for historical and modern buildings (40%-10%). Lo and Ln represent low visual occupancy rates for historical and modern buildings (below 10%). * $p < 0.05$; ** $p < 0.01$.

Table 3. Visual change rate clustering statistics based on k-means.

Engagement	
Cluster1	LnLo-MnLo;MoLn-MoLn;MoLn-LoLn;HoLn-HoLn;HoLn-HoMn;HoMn-HoLn;LoHn-LoMn;LoMn-LoHn;LoHn-LoHn;HoMn-MoMn;MoMn-MoMn;MoMn-MoHn;MoMn-HoLn
Cluster2	MoLn-LoMn;LoMn-LoMn;MoLn-MoMn; HoLn-MoMn

Cluster3	HoLn-MoLn;LoMn-MoMn
Cluster4	MoMn-MoLn;MoLn-HoLn;LoNn-LoMn;MoMn-HoMn;HoLn-HoMn;LoHn-MoMn
Cluster5	HoMn-HoMn
Excitement	
Cluster1	HoLn-MoLn;LoHn-LoMn;HoMn-MoMn;MoMn-HoLn
Cluster2	MoLn-LoMn;LoMn-MoMn
Cluster3	LoMn-LoMn;HoLn-HoLn;MoLn-HoLn;MoMn-HoMn;MoMn-MoHn;HoLn-MoMn
Cluster4	LnLo-MnLo
Cluster5	MoMn-MoLn;MoLn-MoLn;MoLn-MoMn;MoLn-LoLn;HoLn-HoMn;HoMn-HoLn;LoNn-LoMn;LoMn-LoHn;LoHn-LoHn;HoMn-HoMn;MoMn-MoMn; HoLn-HoMn;LoHn-MoMn
Focus	
Cluster1	LoNn-LoMn;HoLn-MoMn;HoLn-HoMn
Cluster2	MoLn-LoMn;HoMn-HoMn
Cluster3	HoLn-HoMn;MoLn-HoLn;HoLn-MoLn;LoMn-LoHn;MoMn-HoMn;LoMn-MoMn
Cluster4	LnLo-Mn;LoHoMn-MoMn;MoMn-MoHn
Cluster5	LoMn-LoMn;MoMn-MoLn;MoLn-MoLn;MoLn-MoMn;MoLn-LoLn;HoLn-HoLn;HoMn-HoLn; LoHn-LoMn;LoHn-LoHn;MoMn-MoMn; LoHn-MoMn;MoMn-HoLn
Interest	
Cluster1	LoMn-LoHn;HoLn-MoMn;HoLn-HoMn
Cluster2	LoMn-LoMn;HoLn-HoMn;HoMn-HoLn; LoHn-MoMn
Cluster3	LoHn-LoMn;HoMn-MoMn;
Cluster4	LnLo-MnLo;MoLn-LoMn;MoMn-MoLn;MoLn-MoLn;MoLn-MoMn;MoLn-LoLn;HoLn-HoLn;MoLn-HoLn;LoNn-LoMn;LoHn-LoHn;MoMn-HoMn;HoMn-HoMn;MoMn-MoMn;MoMn-MoHn;MoMn-HoLn
Cluster5	HoLn-MoLn;LoMn-MoMn
Relaxation	
Cluster1	LnLo-MnLo;MoMn-MoLn;LoMn-LoMn;MoLn-MoLn;MoLn-MoMn;MoLn-LoLn;HoLn-HoLn;HoLn-HoMn; MoLn-HoLn;LoHn-

	LoMn;LoNn-LoMn;LoHn-LoHn;MoMn-MoMn;HoMn-MoMn;HoLn-HoMn;LoMn-MoMn;MoMn-HoLn
Cluster2	LoMn-LoHn
Cluster3	MoLn-LoMn;HoMn-HoLn;HoMn-HoMn
Cluster4	HoLn-MoLn
Cluster5	MoMn-HoMn; MoMn-MoHn;HoLn-MoMn;LoHn-MoMn
Stress	
Cluster1	LnLo-MnLo;HoMn-HoLn;LoHn-LoMn;HoLn-HoMn;LoMn-MoMn
Cluster2	MoLn-LoMn;MoLn-MoLn;MoLn-MoMn;LoNn-LoMn;LoMn-LoHn;LoHn-LoHn;HoMn-MoMn;MoMn-HoLn
Cluster3	LoMn-LoMn;MoLn-LoLn;HoLn-HoLn;HoLn-HoMn; MoLn-HoLn;MoMn-HoMn;MoMn-MoMn;MoMn-MoHn;HoLn-MoMn;LoHn-MoMn
Cluster4	MoMn-MoLn;HoMn-HoMn
Cluster5	HoLn-MoLn

Note: Ho and Hn represent high visual occupancy rates for historical and modern buildings (above 40%). Mo and Mn represent medium visual occupancy rates for historical and modern buildings (40%-10%). Lo and Ln represent low visual occupancy rates for historical and modern buildings (below 10%). "-" indicates a transition from one type to another.

References

Cheng, L., Chu, S., Zong, W., Li, S., Wu, J., & Li, M. (2017). Use of tencent street view imagery for visual perception of streets. *ISPRS International Journal of Geo-Information*, *6*(9), 265. https://doi.org/10.3390/ijgi6090265

Deghati Najd, M., Ismail, N. A., Maulan, S., Mohd Yunos, M. Y., & Dabbagh Niya, M. (2015). Visual preference dimensions of historic urban areas: the determinants for urban heritage conservation. *Habitat International*, *49*, 115–125. https://doi.org/10.1016/j.habitatint.2015.05.003

Foltête, J.-C., Ingensand, J., & Blanc, N. (2020). Coupling crowd-sourced imagery and visibility modelling to identify landscape preferences at the panorama level. *Landscape and Urban Planning*, *197*, 103756. https://doi.org/10.1016/j.landurbplan.2020.103756

Gobster, P. H., Ribe, R. G., & Palmer, J. F. (2019). Themes and trends in visual assessment research: introduction to the landscape and urban planning special collection on the visual assessment of landscapes. *Landscape and Urban Planning*, *191*, 103635. https://doi.org/10.1016/j.landurbplan.2019.103635

Guo, Z., & Sun, L. (2016). The planning, development and management of tourism: the case of dangjia, an ancient village in china. *Tourism Management*, *56*, 52–62. https://doi.org/10.1016/j.tourman.2016.03.017

Jiang, B., Larsen, L., Deal, B., & Sullivan, W. C. (2015). A dose–response curve describing the relationship between tree cover density and landscape preference. *Landscape and Urban Planning*, *139*, 16–25. https://doi.org/10.1016/j.landurbplan.2015.02.018

Li, Z., & Kang, J. (2019). Sensitivity analysis of changes in human physiological indicators observed in soundscapes. *Landscape and Urban Planning*, *190*, 103593. https://doi.org/10.1016/j.landurbplan.2019.103593

Liu, C., & Xu, M. (2021). Characteristics and influencing factors on the hollowing of traditional villages—taking 2645 villages from the chinese traditional village catalogue (batch 5) as an example. *International Journal of Environmental Research and Public Health*, *18*(23), 12759. https://doi.org/10.3390/ijerph182312759

Liu, L., Silva, E. A., Wu, C., & Wang, H. (2017). A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. *Computers, Environment and Urban Systems*, *65*, 113–125. https://doi.org/10.1016/j.compenvurbsys.2017.06.003

Sektani, H. H. J., Khayat, M., Mohammadi, M., & Roders, A. P. (2021). Erbil city built heritage and wellbeing: an assessment of local perceptions using the semantic differential scale. *Sustainability*, *13*(7), 3763. https://doi.org/10.3390/su13073763

Wagtendonk, A. J., & Vermaat, J. E. (2014). Visual perception of cluttering in landscapes: developing a low resolution gis-evaluation method. *Landscape and Urban Planning*, *124*, 85–92. https://doi.org/10.1016/j.landurbplan.2014.01.006

Yang, H., Qiu, L., & Fu, X. (2021). Toward cultural heritage sustainability through participatory planning based on investigation of the value perceptions and preservation attitudes: qing mu chuan, china. *Sustainability*, *13*(3), 1171. https://doi.org/10.3390/su13031171

Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: using machine learning algorithms and google street view imagery. *Applied Geography*, *76*, 147–153. https://doi.org/10.1016/j.apgeog.2016.09.024

Zhang, H., & Lin, S.-H. (2011). Affective appraisal of residents and visual elements in the neighborhood: a case study in an established suburban community. *Landscape and Urban Planning*, *101*(1), 11–21. https://doi.org/10.1016/j.landurbplan.2010.12.010

Zhao, J., Xing, L., Li, R., & Sun, T. (2016). Assessing the visual quality of rural villages in china, compared to urban residential areas. *International Journal of Environmental Studies*, *73*(5), 806–818. https://doi.org/10.1080/00207233.2016.1199418